Stable DNA–protein complexes in yeast *Saccharomyces cerevisiae*. Protein composition analysis

L. Bagdonienė¹, V. Borutinskaitė², R. Navakauskienė², B. Juodka¹

¹ Department of Biochemistry and Biophysics, Vilnius University, M. K. Ciurlionio 22, LT-03101 Vilnius, Lithuania
² Department of Development Biology, Institute of Biochemistry, Mokslininkų 12, LT-08662 Vilnius, Lithuania

Proteins tightly bound to DNA (TBP) are a protein group that remains attached to DNA by covalent or non-covalent bonds after its deproteinisation [1, 2]. These proteins are released only after DNA digestion with DNase I or benzonase. TBP have been found in DNA of numerous evolutionary distant species [3, 4]. Only a few of these TBP have been characterized. In the present work, we have identified some proteins that are present in *Saccharomyces cerevisiae* TBP–DNA protein complexes by two-dimensional electrophoresis and MALDI-TOF MS analysis. Most of these proteins (chromatin assembly factor 1, NNF1 protein, DNA damage tolerance protein RHC31, DNA repair protein RAD 7, SOH1 protein, etc.) have been found to participate in chromatin reorganization and regulation processes.

Key words: DNA–protein interaction, stable DNA–protein complexes

Abbreviations: TBP – eukaryotic non-histone proteins tightly bound to DNA.

INTRODUCTION

Like in higher eukaryotes, the organization of DNA into chromatin and chromosomal structures plays a central role in many aspects of yeast cell biology. The processes ranging from chromosome stability and segregation to gene expression are intimately linked to chromatin configuration [5, 6]. What proteins are involved in the 3-dimensional genome organization is not completely clear. The polypeptides that are able to form permanent or transient tightly bound or covalent complexes with DNA are of special interest [7–10]. A small but persistent amount of nuclear polypeptides remains associated with eukaryotic DNA submitted to the most rigorous deproteinization methods, including treatment with SDS, proteinase K and partially with phenol. Usually, proteins can easily be removed from DNA with high ionic strength buffers and with dissociating agents such as urea, guanidine HCl, SDS, although even after treatment with such reagents there are some proteins (TBP) that remain tightly bound with DNA [7–10]. They are released from DNA only after DNA digestion with DNase I or benzonase. TBP distribution in the genome is site-specific. They are enriched in sequences of nuclear matrix attachment sites. These proteins are part of insoluble nuclear matrix proteins [11–13]. Only a few of these proteins have been characterized. Recently it has been demonstrated that some of such proteins, tightly bound to DNA separated from Erlich ascite and yeast *Saccharomyces cerevisiae* cells, show phosphatase and kinase activity [14]. It was shown that a fraction of three serpins, Spi-1, Spi-2 and Spi-3, has the capability to constitute a proteinase K / SDS stable DNA–protein complex [7]. The stability of these proteins to proteinase K action may be associated with the protective function of DNA and with the presence of protease inhibitors (Spi-1, Spi-2 and Spi-3) in these complexes. C1D was the first expressed 16 kDa tightly bound to DNA protein. C1D functions as a component of a complex involved in the transcription repression [8]. However, the functions and regulatory mechanism of TBP-DNA proteins is not fully understood.

MATERIALS AND METHODS

Cell lysis and salting-out procedure [15]. Cells were disrupted in liquid nitrogen and suspended (1 : 6, v/v) in buffer (10 mM Tris-HCl, pH 8.2, 400 mM NaCl, 2 mM EDTA). Proteinase K (Sigma), dissolved in 1% SDS, was added to the final concentration of 50 μg/ml. Cell lysis was induced by addition of 0.2 vol 10% SDS per 3 vol of cell suspension. The lysis mixture was sheared by pressing it through a 1.5 mm gauge needle and incubated for 18 h at 37 °C. Then the lysates were diluted (1 : 3, v/v) with a suspension buffer containing 0.63% SDS, mixed with 6 M NaCl solution (3 : 1 v/v) and chilled for 1 h before centrifugation (6000 g, 4 °C). Such DNA crude preparations were precipitated
with ethanol (1 : 2 v/v) and then digested with pancreatic ribonuclease and DNase I as in [4].

**Chromatography on Heparin-sepharose Fast Flow column.** 1 mg TBP proteins were applied onto a heparin sepharose column (0.5 × 9 cm). The column was washed with 25 ml solution (50 mM tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT, 5 mM MgCl₂, 10 mM β-mercaptoethanol, 0.5 mM EGTA, 0.4 mM PMSF), then with 70 / 70 ml gradient of 0–1 M NaCl in 50 mM tris-HCl pH 7.5 containing 1 mM EDTA, 1 mM DTT, 5 mM MgCl₂, 10 mM β-mercaptoethanol, 0.5 mM EGTA, 0.4 mM PMSF. Flow rate 0.3 ml/min. The concentration of the fractions was performed with Polyethylenglycol 20000. Fraction 2, eluted with 0.05 M NaCl solution (Fig. 1), was applied on 2DE gel electrophoresis.

**SDS-PAGE gel electrophoresis.** 12% SDS-PAGE gels were run at 120 V constant voltage for 4 hours [16]. Gels were stained with Coomassie dye.

**2DE gel electrophoresis (IEF / SDS-PAGE).** The samples were dissolved in a sample buffer (12 g urea, 50 mg DTT, 0.13 ml pharmalit 3–10, 0.13 ml TritonX-100, water until 25 ml). For the first dimension, Immobiline Dry Strips 4–7 (13 cm) were used (Amersham Biosciences). Dry strips were rehydrated, reduced and alycylated according to manufacturers’ recommendations. For the second dimension, it 8–18% Excell gels (Amersham Biosciences) were used. IEF / SDS-PAGE were performed with a Multiphor II device (Amersham Biosciences). SDS buffer strips: anodal (0.45 mol/l Tris-acetate, pH 6.6, 4 g/l SDS, 0.05 g/l Orange G), cathodal (0.08 mol/l Tris, 0.80 mol/l Tricine and 6 g/l SDS, pH 7.1). Gels were visualized with Coomasie dye.

**Protein determination.** Protein content in the chromatography steps during purification procedure was followed by measuring the absorbance at 280 nm. In the pooled samples, protein concentration was determined by the method of [17], using BSA as a standard.

**In-gel tryptic digestion and MALDI-TOF MS.** The areas of the gel that were deemed to be of interest were cut out and subjected to in-gel tryptic digestion overnight [18], the gel slides were dehydrated with 50% acetonitrile and then dried completely using a centrifugal evaporator (DNA Mini, Epperdorf). The protein spot was rehydrated in 20 µl of 25 mM ammonium bicarbonate (pH 8.3) containing 20 µg/ml modified trypsin (Promega). Once this solution was fully absorbed by the gel, a trypsin-free buffer was added just enough to cover the slice, and the samples were incubated overnight at 37 °C. The tryptic peptides were subsequently extracted from the gel slides as follows. Any extraneous solution remaining after the digestion was removed and placed in a fresh tube. The gel slides were first subjected to an aqueous extraction and then to organic extraction with 5% trifluoroacetic acid in 50% acetonitrile, shaking occasionally. The digestion and extract solutions were then combined and evaporated to dryness. For MALDI-TOF (Matrix-Assisted Laser Desorption / Ionization Time-Of-Flight) analysis, the peptides were redissolved in 3 µl of 30% acetonitrile and 0.01% trifluoroacetic acid and were then prepared with a matrix (α-cyano-4-hydroxicinnamic acid) on the target plate. The analysis was performed on a Voyager MALDI-TOF MS (Perspective Biosystems Inc., Town State) and externally calibrated using synthetic peptides with known masses. The spectra were obtained in the positive ionization mode at 20 kV. The mass information generated from the composite spectrum was submitted to a search performed with the Protana or EXPASY database, using the MS-Fit algorithm or the PeptIdent search engine.

**RESULTS AND DISCUSSION**

Recently it has been demonstrated that DNA obtained from *Saccharomyces cerevisiae* by SDS and proteinase K treatment still contains polypeptides which are resistant to SDS, urea, etc. It seems that DNA–protein interaction in these complexes is extremely stable, and in some cases the DNA–protein linkage is possibly covalent [4, 14].

![Fig. 1. Yeast TBP fractionation by heparin–sepharose chromatography:](image)

**A.** 1 mg total protein was applied onto a heparin-sepharose column (0.5 × 9 cm). Column washed with 25 ml solution (50 mM tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT, 5 mM MgCl₂, 10 mM β-mercaptoethanol, 0.5 mM EGTA, 0.4 mM PMSF), then with 70 / 70 ml gradient of 0–1 M NaCl in the same buffer. Flow rate 0.3 ml/min;

**B.** 10% SDS-PAGE stained with Coomassie dye: 1–4 – fractions after heparin-sepharose chromatography, 5 – TBP mixture obtained after digestion of DNA with benzonase
This study aims to further characterize tightly DNA-bound complexes of proteins from *Saccharomyces cerevisiae* by two-dimension electrophoresis and MALDI-TOF MS analysis.

Yeast DNA was isolated according to Miller’s method [15] by cell lysis in the presence of SDS and proteinase K (50 µg/ml), followed by salting out of residual proteins with NaCl (1.6 M). Some proteins co-purified with DNA remain tightly bound with DNA after treatment with reagents such us urea, guanidine HCl, SDS, mercaptoethanol, phenol. After DNA hydrolysis with DNase I, SDS-PAGE revealed a high molecular weight or aggregated proteins – 70–40 kDa and 20–15 kDa [14].

These proteins behave as complex or aggregated proteins. They are difficult to separate by electrophoresis or chromatography. Application of DE sepharose, SP sepharose, CE sepharose for *Saccharomyces cerevisiae* TBP chromatography was not successful. Proteins were eluted from columns as one group in the same fraction (results not shown). Only fractionation by heparin sepharose was a little better (Fig. 1). For analysis of yeast TBP by two-dimensional electrophoresis, a TBP mixture obtained after DNA digestion by benzonase and protein fraction 2 obtained after fractionation of this TBP protein mixture by heparin sepharose was taken.

Most of TBP separated by 2 DE are acid proteins, pl 4.2–6 (Fig. 2A, B). Some of these protein bands were cut from the 2DE gels digested with trypsin and analysed with MALDI-TOF MS. Tryptic digestion of proteins gives peptides of different length depending on the distribution of the corresponding trypptic cleavage sites (arginine and lysine).

Protein identification was performed by searching in a protein sequence database (SwisProt) using the PeptIdent and MS-Fit programs. The following parameters were used for database searches: monoisotopic mass accuracy 50–100 ppm, missed cleavages 0–1 and complete carbamidomethylation of cysteines. The list of proteins identified employing the PeptIdent program is presented in Table.

Most proteins identified by the MALDI-TOF MS method are somehow associated with DNA functions. Among them, we found no known chromatin proteins such as histones or soluble proteins.

### Table

<table>
<thead>
<tr>
<th>Band No</th>
<th>pI</th>
<th>Mw (kDa)</th>
<th>Peptident</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Calculated</td>
<td>Experimental</td>
<td>Score</td>
</tr>
<tr>
<td>1</td>
<td>5.92</td>
<td>51.25</td>
<td>51</td>
<td>0.04</td>
</tr>
<tr>
<td>2</td>
<td>4.86</td>
<td>82.00</td>
<td>52</td>
<td>0.33</td>
</tr>
<tr>
<td>3</td>
<td>4.93</td>
<td>48.25</td>
<td>48</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>5.38</td>
<td>47.96</td>
<td>46</td>
<td>0.4</td>
</tr>
<tr>
<td>5</td>
<td>4.48</td>
<td>35.96</td>
<td>43</td>
<td>0.07</td>
</tr>
<tr>
<td>6</td>
<td>4.89</td>
<td>23.63</td>
<td>22</td>
<td>0.13</td>
</tr>
<tr>
<td>7</td>
<td>4.91</td>
<td>16.07</td>
<td>16</td>
<td>0.33</td>
</tr>
<tr>
<td>8</td>
<td>5.2</td>
<td>39.27</td>
<td>39</td>
<td>0.06</td>
</tr>
<tr>
<td>9</td>
<td>4.55</td>
<td>70.95</td>
<td>70</td>
<td>0.05</td>
</tr>
<tr>
<td>10</td>
<td>5.15</td>
<td>49.39</td>
<td>27</td>
<td>0.12</td>
</tr>
<tr>
<td>11</td>
<td>4.45</td>
<td>14.74</td>
<td>16</td>
<td>0.1</td>
</tr>
<tr>
<td>12</td>
<td>5.33</td>
<td>63.8</td>
<td>50</td>
<td>0.05</td>
</tr>
</tbody>
</table>

1st column — the number of band in gels (Fig. 2A, B); 2nd column — Mw and pI proteins calculated from the amino acid sequence in ExPasy database; 3rd column — Mw calculated according to their migration in 2 DE; 4th column — MOWSE scores and sequence coverage (%); 5th column — the list of identified proteins.
nuclear enzymes (Dnases, Rnases, RNA, DNA polimerases, etc.) However, all of these TBP proteins seem to participate in chromatin rearrangement and regulation processes.

We have identified that 45 kDa protein (Fig. 2A, band No. 1) is a Saccharomyces cerevisiae chromatin assembly factor 1 P60 subunit (CAF-1 60 kDa subunit). This protein functions in chromatin assembly during DNA replication and repair and plays a role in the maintenance of genome stability. DNA replication and repair require a new chromatin assembly during which the four core histones must be assembled into a histone octamer containing two H2A_H2B dimers and a H3_H4 tetramer wrapped in DNA to form a nucleosome. Many histone chaperones are known to be involved in nucleosome and chromatin assembly, of which chromatin assembly factor 1 (CAF-1) is among those best most studied [19]. In Saccharomyces cerevisiae, deletion of CAF-I leads to an increased UV sensitivity and silencing defects at telomeres as well as at the mating type loci [20]. CAF-I defects also might result in activation of the DNA-damage checkpoint [19].

Using MALDI-TOF MS analysis, we have also identified ARS binding protein 1 (Protein ABF1) in the 52 kDa protein band (Fig. 2A, band No. 2). Protein ABF1 is a sequence-specific DNA-binding protein with possible chromatin reorganizing activity involved in transcription activation, gene silencing, DNA replication and repair. It has been indicated that ABF1 interacts with RAD 16, RAD 7 proteins which recognize and bind damaged DNA in an ATP-dependent manner [21]. ABF1 is extensively phosphorylated on Ser and Thr residues and binds to the 5'-TCN(7)ACG-3' sequence in many promoters [22].

Proteins in bands No. 3 (43 kDa) and 4 (35 kDa) (Fig. 2A) were identified as a 26S protease regulatory subunit 6A homolog (TBP-1) and a regulatory subunit 6B homolog (YNT1 protein), (TAT-binding homolog 2), respectively. These proteins are 26S proteasome AAA--ATPase subunits. Proteasome proteins are a well-established component of the nuclear matrix [23]. Proteasomes are involved in the degradation of numerous transcription factors and other nuclear proteins, including the tightly bound C1D protein [8]. A growing body of evidence demonstrates that the components of the proteasome are involved directly and mechanistically in the regulation of gene transcription. For transcription to be initiated appropriately, hundreds of polypeptides must interact with DNA at the right place, at the right time, and in response to the right signals. Orchestrating all of these interactions requires that massive protein complexes be escorted on and off the chromatin, and that signals unique to each phase of the process – initiation, promoter clearance, elongation – be terminated when appropriate. Exactly how the cell accomplishes this task is not clear, but emerging evidence suggests that one way is through recruitment of protein machine – the proteasome. Proteolytic activities of the proteasome are important for establishing the limits for transcription, for promoting the exchange of transcription factors on chromat, and possibly for facilitating multiple rounds of transcription initiation. Non-proteolytic activities of the proteasome are important for co-activator recruitment, transcriptional elongation, and histone modification. Emerging evidence in yeast suggests roles for ATPases of the 19s proteasome as mediators of transcriptional systems through their association with actively transcribed promoters, facilitation of clearance of post-elongation complexes and recruitment of coactivators [24–26]. The mechanisms of action and independent function of the 19S ATPases on promoters are unclear and remain to be fully addressed. Observations in yeast that 19S ATPases link histone ubiquitination and histone methylation [27] and that the 19S recruits histone-modifying enzymes to promoters [28] suggest that the 19S may independently regulate transcription by controlling chromatin modifications [29].

We also identified another three proteins involved in ubiquitin-related processes. 30 kDa protein in band No. 5 (Fig. 2B) is identified as Saccharomyces cerevisiae DOS2 protein. This protein acts in ubiquitin metabolism and is necessary for the control of single-copy DNA replication. In eukaryotic cells, DNA replication is initiated at many sites, often following a complex temporal program. The loss of ubiquitin hydrolase could affect DNA replication in at least two ways. First, a general slowing of proteolysis due to proteasome inhibition could cause persistence of the factor that promotes DNA replication. Second, if the activity of a the factor that promotes or limits replication is controlled by reversible modification with ubiquitin, the absence of ubiquitin hydrolase could alter the state of this regulation [30].

The 39 kDa protein band No. 8 (Fig. 2B) is a DNA damage tolerance protein RHC31, RAD31 homolog. This protein could be involved in a ubiquitin-related process important for DNA damage tolerance [31, 32].

The 30 kDa protein in band No. 12 (Fig. 2B) was identified as a DNA repair protein RAD 7, involved in excision repair of DNA damaged with UV light, bulky adducts, or cross-linking agents, and possessing DNA-dependent ATPase activity. It participates in protein ubiquitination during the ubiquitin-dependent protein catabolic process [33]. Proteins RAD31 and RAD 7 are components of the same DNA repair protein complex. It was shown also that protein RAD 7 interacts with protein ABF1 [21] found in Saccharomyces cerevisiae TBP complex (Fig. 2A, band No. 2).

Using the MALDI-TOF MS analysis, we have also identified two proteins that belong to RNA polymerase II mediator complex. 16 kDa protein in band No. 7 (Fig. 2B) shows a homology with the suppressor of RNA polymerase II SBR7 / MED21, and 16 kDa protein in band No. 11 shows a homology to the suppressor of RNA polymerase II, SOH1 / MED31. These proteins belong to the mediator middle submodule [34–36]. The RNA polymerase II mediator complex is composed of at least 21 subunits that form three structurally distinct sub-modules. The mediator head module contains MED6, MED8, MED11, SRB4 / MED17, SRB5 / MED18, ROX3 / MED19, SRB2 / MED20 and SRB6 / MED22, the middle module contains MED1, MED4, NUT1 / MED5, MED7, CSE2 / MED9, NUT2 / MED10,SRB7 / MED21 and SOH1 / MED31, while the tail module contains MED2, PGD1 / MED3, RGR1 / MED14, GAL11 / MED15 and SIN4 / MED16. The mediator functions as a bridge to convey information from gene-specific regulatory proteins to the basal RNA polymerase II transcription machinery. The mediator complex, having a compact confor-
mation in its free form, is recruited to promoters by direct interactions with regulatory proteins and serves for the assembly of a preinitiation complex with RNA polymerase II and the general transcription factors.

In stable DNA–protein complexes, we have also identified the chromatin modification-related protein EAF7 (band No. 10, Fig. 2B). This protein is a component of the NuA4 histone acetyltransferase complex which is involved in transcriptional activation of selected genes, principally by acetylation of the nucleosomal histone H4 and H2A. The NuA4 complex is also involved in DNA repair [37].

22 kDa protein in band No. 6 (Fig. 2B) shows a homology with *Saccharomyces cerevisiae* NNF1 protein. Nnf1p is a protein associated with the nuclear envelope in *Saccharomyces cerevisiae* and is required for a number of nuclear functions. Cells depleted of nnf1 or containing a temperature-sensitive NNF1 mutation have elongated microtubules and become bi- and multinucleate. They also have a fragmented nucleolus and accumulate poly(A) + RNA in the nucleus [38, 39].

It has been demonstrated earlier that some of tightly DNA-bound proteins separated from EAT and *Saccharomyces cerevisiae* cells exhibit phosphatase and kinase enzymatic activities [4, 14]. The MALDI-TOF MS analysis revealed that 43 kDa protein in band No. 9 (Fig. 2A) is a protein phosphatase PP2A regulatory subunit A (PR65). It is known that the PR65 / A subunit of protein phosphatase 2A serves as a scaffolding molecule to coordinate the assembly of the catalytic subunit and a variable regulatory B subunit of phosphatase PP2A, generating functionally diverse heterotrimeric [40]. The proposed targets of PP2A activity include cohesin and Sgo, both of which would otherwise dissociate from chromosomes upon phosphorylation by kinase [39].

Another protein (35 kDa) in band No. 13 (Fig. 2B) maches *Saccharomyces cerevisiae* serine / threonine-protein kinase CHK1 (Checkpoint kinase 1). This protein is involved in cell cycle arrest when DNA damage has occurred or when unligated DNA is present. [42]. Activated CHK1 has a full spectrum of substrates that are key cell cycle regulators. For example, CHK1-associated p53 phosphorylation leads to stabilization of the tumor suppressor and causes subsequent cell cycle arrest and apoptosis. In contrast, in response to DNA damage, CDC25A, a positive regulator of cyclin-dependent kinases, is hyperphosphorylated by CHK1 and degraded rapidly through an ubiquitin-mediated pathway to block cell cycle progression [43].

Mass spectrometry analysis enabled us to characterize a set of TBP in the yeast *Saccharomyces cerevisiae*. Yeast TBPs manifest homologies to functionally different enzymes and regulatory factors that participate in chromatin modification, reconstruction and repair. These proteins tightly interact with DNA and are important for the specific structural organization of DNA and proteins in the nucleus. TBPs seem to perform specific chromosomal functions, but they are also versatile molecules participating in multisubunit complexes. At present, we know many examples of proteins which participate in different processes and are components of different complexes. We have no answer why these TBPs are so strongly associated with DNA. A more detailed characterization of TBP structure and their localization in cells in vivo can shed some light on the biochemical and biological activities that are not presently understood.

This work was supported by the Lithuanian State Science and Studies Foundation grant No. T-109/07.

**References**


3. **Tsanev R., Avramova Z. Chromosoma 1994; 103: 293–301.**

4. **Labeikyte D, Bagdoniene L, Juodka B. Biologija 1999; 4: 22–7.**


22. **Francesconi SC, Eisenberg S. PNAS 1991; 88: 4089–93.**


27. **Ezhkova E, Tansey WP. Mol Biol Cell 2004; 14: 345–42.**


Baltymai (TBP), išskirti iš įvairaus evoliucijos lygmens organizmų, pasižymi ypač tvirta sąveika su DNR: jų neįmanoma pašalinti nuo DNR detergentais, karbamidu, redukuojančiais agentais. Šiame darbe apibūdinami tvirtai su DNR susieti *Saccharomyces cerevisiae* mielių baltymai. Mielių DNR buvo išskirta Miller metodu: iš pradžių ląstelės suardomos NDS ir proteinaze K (50 µg/ml), vėliau baltymai nusodinami 1,6 M NaCl. Paveikus DNR baltyminius kompleksus benzonaze, baltymai buvo frakcionuoti naudojant heparino sefarozės kolonėlę ir 2DE elektroforezę. Baltymų sudėtis nustatyta MALDI-TOF MS metodu. Visi identifikuoti nukleoproteidinio komplekso baltymai (su chromatinu asocijuotas veiksnys 1, NNF1 baltymas, DNR pažaidų tolerancijos baltymas RHC31, DNR reparacijos baltymas RAD 7, SOH1 baltymas ir kiti) yra susiję su DNR funkcijomis ir yra svarbūs chromatino pertvarkymui, galbūt DNR pritvirtinimui prie branduolio matrikso. Gauti mielių DNR baltyminių kompleksų baltymų proteominės analizės rezultatai kol kas neleidžia pasakyti, ar šie baltymai yra funkciškai susiję tarpusavyje ir veikia *in vivo* kaip vieno komplekso dalis. Visi jie vienaip ar kitaip gali reguliuoti chromatino veiklą.